AD_____

Award Number: DAMD17-96-1-6228

TITLE: Demonstration Project on Mammographic Computer-Aided
Diagnosis for Breast Cancer Detection

PRINCIPAL INVESTIGATOR: Kunio Doi, Ph.D.

CONTRACTING ORGANIZATION: The University of Chicago
Chicago, Illinois 60637

REPORT DATE: October 2002

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are
those of the author(s) and should not be construed as an official
Department of the Army position, policy or decision unless so
designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 2002	Final (26 Sep 96 – 25 Sep 02)

4. TITLE AND SUBTITLE

Demonstration Project on Mammographic Computer-Aided Diagnosis for Breast Cancer Detection

5. FUNDING NUMBERS

DAMD17-96-1-6228

6. AUTHOR(S)

Kunio Doi, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The University of Chicago
Chicago, Illinois   60637

E-Mail: k-doi@uchicago.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland   21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

The goal of this project is to demonstrate the clinical usefulness of computer-aided diagnosis (CAD) in mammographic detection of breast cancer. Our plan is to develop advanced CAD schemes for detection and characterization of clustered microcalcifications and masses by incorporating artificial neural networks and various image processing techniques. Clinical mammography workstations for automated detection of suspicious lesions in mammograms will be developed by integration of laser digitizer, high-speed computer and advanced CAD software. The prototype workstations will be used as a "second opinion" in interpreting mammograms by reducing observational errors. The outcomes of radiologists' image readings in the detection of breast cancer will be evaluated by examining radiologists' performance when reading films only and when reading film with the computer results. We believe that the outcomes of this demonstration project will lead to large-scale clinical trials and will result in commercial projects for practical routine use in breast imaging.

14. SUBJECT TERMS
Breast Cancer, CAD Schemes, CAD Software, Demonstration Project

15. NUMBER OF PAGES
58

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

1

1. INTRODUCTION

This final report is prepared at the end of the sixth year of a 6-year project (originally a 4-year project with two year extension to 25 October 2002). Therefore, the sections for 1.1 Subject and Scope of the Research, 1.2 Purpose, 1.3 Background of Previous Work, and 2.1 Experimental Methods, Assumptions and Procedures (pages 2-16) are the same as those in the last year's report. In the subsequent sections, some parts in the last year's report are kept the same for continuation.

1.1. The Subject and Scope of the Research

Breast cancer is a leading cause of death in women, with an estimated 46,000 deaths per year in the United States (ref. 1). Mammography is currently the only known reliable method for early detection of breast cancer (refs. 2,3). However, early mammographic signs of breast cancer such as clustered microcalcifications and masses are usually very subtle, and thus 10-30% of lesions are missed even by trained radiologists. These misses are due to the often low conspicuity of lesions, eye fatigue, and human error (refs. 4-7). However, there is clear evidence (refs. 8,9) that radiologists' accuracy in the detection of subtle breast lesions would be improved if a computer output indicating possible sites of suspicious lesions were made available to radiologists as a "second opinion."

As a team of investigators at the Kurt Rossmann Laboratories for Radiologic Image Research at the University of Chicago, we have been involved since 1985 in developing the concepts and methodology of computer-aided diagnosis (CAD) with which to assist radiologists in detecting lesions and improving the sensitivity of breast cancer diagnosis through mammography (refs. 10, 11). CAD may be defined as a diagnosis made by a radiologist who takes into account the results of automated computer analyses of radiographic images. The computer output may be used as a "second opinion." We have extensive experience in developing CAD schemes. In addition to breast cancer, we have developed computer schemes for the detection of lung nodules

2

(refs. 12, 13), interstitial infiltrates (refs. 14, 15), cardiomegaly (refs. 16, 17) and pneumothorax (ref. 18) in chest radiography; the detection of stenotic lesions and blood flow analysis in angiography (refs. 19, 20); and the assessment of osteoporosis in skeletal radiography (ref. 21).

In mammography, CAD schemes are being developed for detection of clustered microcalcifications (refs. 8, 24-26, 28) and for detection of masses (refs. 22, 23). A basic scheme for automated detection of clustered microcalcifications employs a difference image technique to enhance the signal-to-noise ratio of microcalcifications, followed by thresholding, feature extraction and classification using artificial neural networks. At present, the performance of this scheme provides a sensitivity of approximately 85% in the detection of clustered microcalcifications with a false positive rate of approximately 0.7 per mammogram when it is tested on our database of 78 mammograms, in which one half are normal cases and the other half includes subtle clustered microcalcifications.

For the automated detection of mammographic masses, another CAD scheme is being developed on the basis of a bilateral subtraction technique that analyzes deviations of architectural symmetry between the right and left breast images, with asymmetries indicating potential masses (refs. 29-32). Currently, this scheme performs at approximately 90% sensitivity with a false positive rate of about 2 per case when it is tested on our database of 154 pairs of mammograms. Our current research effort on these CAD schemes is focused primarily on improving further their performance through careful analysis of computer false-positives and false-negatives.

To date, these studies have been performed retrospectively on selected sets of mammograms, and we have obtained results that indicate that our schemes have the potential to be used as an effective aid for radiologists. We are now at the stage in the development of our CAD program to test our schemes prospectively on a large number of clinical mammograms.

3

On November 8th, 1994, we implemented an "intelligent" mammography workstation (ref. 34) and began the first test of our schemes on clinical screening mammograms obtained in the mammography section of our department. This workstation consists of an IBM Powerstation 590, a Konica LD4500 laser film digitizer, an Alphatronix Inspire 40-GB magneto-optical jukebox, two Imlogix 1024 line monitors, and a Seikosha VP 4500 video printer for hard copy. The "intelligence" of the workstation comes from our automated detection schemes for clustered microcalcifications and masses.

In order to realize clinically and practically mammographic CAD for detection of breast cancers in screening programs, it is necessary to have commercial products for widespread use by radiologists in breast clinics, community hospitals, and academic medical centers. Therefore, in 1993, ARCH Development Corporation (ARCH), which is a not-for-profit organization created by the Board of Trustees of the University of Chicago in 1986 as a unique mechanism to commercialize inventions developed by the faculty at the University of Chicago and by scientists at Argonne National Laboratory, licensed its inventions on CAD and related technologies to R2 Technology, Inc.

R2 Technology, Inc. was founded in 1993 with the specific goal of developing and marketing a computer aided diagnostic system in mammographic detection of breast cancer. R2 Technology, Inc., has been funded by leading venture firms in Silicon Valley-Sigma Partners, and Burr, Egan, Deleage and Co. -- that have supported many other successful medical and computer companies. Its development group has over 250 man-years of experience in medical imaging and computer systems. Its product development process has been to identify those high potential prototype systems in leading research institutions and form alliances and integrate those systems into R2's core technology. As of May 1995, R2 has alliances with the University of Chicago, Lockheed Missiles & Space Company, Inc., and Sandia National Laboratories.

Therefore, the next logical step in the development of CAD is to conduct a large-scale, multi-institutional demonstration project to examine whether additional breast

4

cancers can be found by use of mammographic CAD workstations. We believe that the performance of mammographic CAD schemes has reached the high level necessary for clinical evaluation. Serious efforts toward commercialization of CAD units have already begun. Therefore, it is likely that a clear positive outcome from this study would result in production of commercial products for widespread use in breast imaging and would lead to large-scale clinical trials.

1.2. Purpose

The goal of this project is to demonstrate the clinical usefulness of computer-aided diagnosis (CAD) in mammographic detection of breast cancer through multi-disciplinary and multi-institutional efforts. We plan to develop clinical prototype mammography workstations for automated detection of suspicious lesions in mammograms by incorporating image processing techniques and artificial neural networks. The prototype workstation will be used as a "second opinion" to assist radiologists' interpretation of mammograms. Clinical usefulness of the mammography CAD will be demonstrated and evaluated at four hospitals in the Chicago area. The major hypothesis to be tested in this proposal is that CAD improves accuracy in the detection of breast cancer by reducing observational errors on mammographic images. Our proposal is designed to demonstrate that approximately 23 additional breast cancers will be detected among approximately 45,000 screenees due to the use of CAD computer output.

The specific aims of this demonstration project are listed below.

(1) Further development of advanced CAD schemes for detection of breast lesions

   (a) Automated detection scheme for clustered microcalcifications

   (b) Automated detection scheme for masses

   (c) Automated scheme for characterization of detected breast lesions

(2)   Development of the prototype mammography CAD workstations by integration of

laser digitizer, high-speed computer, and advanced CAD schemes

(3)   Clinical demonstration and evaluation of prototype mammography CAD

workstations at two hospitals: one academic institution and one community hospital

(4)   Analysis of outcomes of the clinical evaluation of the prototype workstations for

detection of additional breast cancers by the use of computer output

1.3. Background of Previous Work

We have been working on the development of computer-aided diagnostic (CAD)

schemes for mammography, chest radiography, angiography, and bone radiography since

1985. Therefore, we have extensive experience in quantitative analysis of radiographic

images for detection and characterization of various patterns based on computer-vision

methods and artificial neural networks. These extensive studies provide the basis for the

continued development and testing of advanced CAD schemes for the detection of breast

lesions proposed in this research. A number of investigations which are relevant to this

study are described briefly here.

(1)   Development of computerized detection scheme for mammographic

microcalcifications

We have investigated the application of computer-based methods to the detection

of microcalcifications on digital mammograms. Our computer detection system was

based on a difference-image technique in which a signal-suppressed image was

subtracted from a signal-enhanced image to remove the structured background in a

mammogram (ref. 24). Signal-extraction techniques adapted to the known physical

characteristics of microcalcifications were then used to isolate microcalcifications from

the remaining noise background (ref. 25). Signal-extraction criteria based on the size,

contrast, number, texture, and clustering properties of microcalcifications were next

imposed on the detected signals to distinguish true signals from noise or artifacts (refs. 8,

25). The detection accuracy of the computer scheme was evaluated by means of a free-response receiver operating characteristic (FROC) analysis. In a study of 78 clinical images containing subtle microcalcifications, the automated computer scheme obtained an 85% true-positive cluster detection rate at a false-positive detection rate of 1.5 clusters per image. These results indicated that the automated method has the potential to aid radiologists in screening mammograms for clustered microcalcifications.

We have applied a shift-invariant neural network (SIANN) to eliminate false-positive detections reported by the CAD scheme. The SIANN is a layered feed-forward neural network with local, spatially-invariant interconnections (refs. 27, 28). The basic idea of local, spatially-invariant interconnections (or sharing local interconnection weights) was first introduced by Fukushima in his Neocognitron for recognition of handwriting characters in the early 1980s (refs. 35, 36). The SIANN developed by Zhang et al. (ref. 27) for image processing is a feed-forward neural network without the lateral interconnections and feedback loops that are included in the Neocognitron. Furthermore, a modified error backpropagation (EBP) algorithm with the shift-invariant-connection constraint (ref. 27) is used as the training algorithm in the SIANN. The SIANN has been shown to be a powerful tool for pattern recognition and image processing, since it can learn to discriminate between objects on the basis of local features with results that are invariant to translation of the objects (refs. 27, 28).

This neural network was trained to detect each individual microcalcification in a given region of interest (ROI) reported by the CAD scheme. A ROI was classified as a positive ROI if the total number of microcalcifications detected in the ROI was greater than two. The performance of the shift-invariant neural network was evaluated by means of a jack-knife method and conventional receiver operating characteristic (ROC) analysis by using a database of 168 ROIs that had been reported by the CAD scheme when applied to 39 mammograms. The analysis yielded an average area under the ROC curve (A_z) of 0.91. Approximately 55% of false-positive ROIs were eliminated without any

loss of true-positive ROIs (ref. 28). This result was considerably better than that obtained in our previous study using a conventional three-layer, feed-forward neural network.

We have also studied radiologists' detection of clustered microcalcifications on mammograms to determine whether CAD can improve radiologists' performance. The results of a ROC study showed that CAD, using the level of computer performance at that time (sensitivity = 87%, 4 false clusters per image), does significantly ($p<0.001$) improve radiologists' accuracy in detecting clustered microcalcifications under conditions that simulate the rapid interpretation of screening mammograms (ref. 8). The results also suggested that a reduction in the computer's false-positive rate would further improve radiologists' diagnostic accuracy.

The importance of our findings is that a computerized scheme can detect clustered microcalcifications in digitized mammograms at a high level of sensitivity that would be comparable to levels obtained by radiologists. In addition, radiologists' performance in the detection of clustered microcalcifications can be improved significantly when the results of the computer output are provided as an aid to the radiologists.

(2)   Development of computerized detection schemes for mammographic masses

A computerized scheme has been developed for the detection of masses in digital mammograms. Based on deviations from the normal architectural symmetry of the right and left breasts, a bilateral subtraction technique was used to enhance the conspicuity of possible masses. The scheme employed pairs of conventional screen-film mammograms (right and left MLO views and right and left CC views), which were digitized by a TV camera/Gould digitizer. The right and left breast images in each pair were aligned manually during digitization. A nonlinear bilateral subtraction technique, which involves linking multiple subtracted images, was investigated and compared to a simple linear subtraction method (refs. 29, 30). Various feature-extraction techniques were used to reduce false-positive detections resulting from the bilateral subtraction. The scheme was

8

evaluated using 46 pairs of clinical mammograms and was found to yield a 95% true-positive rate at an average of three false-positive detections per image. This preliminary study indicated that the scheme is potentially useful as an aid to radiologists in the interpretation of screening mammograms.

We continued to investigate the characteristics of actual masses and non-mass detections in order to develop feature-analysis techniques with which to reduce the number of non-mass (i.e., false-positive) detections. These feature-analysis techniques involved extraction of various features such as area, contrast, circularity and border-distance based on the density and geometric information of masses in both processed and original breast images. Cumulative histograms of both actual-mass detections and non-mass detections were used to characterize extracted features and to determine the cutoff values used in the feature tests. The effectiveness of the feature-analysis techniques was evaluated using FROC analysis. Results showed that the feature-analysis techniques effectively improved the performance of the computerized detection scheme: about 35% of false-positive detections were eliminated without loss in sensitivity (ref. 31).

We have developed an automated technique for the alignment of right and left breast images for use in the computerized analysis of bilateral breast images. In this technique (ref. 32), the breast region was first identified by use of histogram analysis and morphological operations. The anterior portions of the tracked breast border and computer-identified nipple positions were selected as landmarks for image registration. The paired right and left breast images were then registered relative to each other by use of a least-squares matching method. Based on FROC and regression analyses, the detection performance obtained with the automated alignment technique was found to be higher than that obtained with simulated misalignments. These results indicated that automatic alignment of breast images is feasible and that mass-detection performance appears to improve with the inclusion of asymmetric anatomic information and is not sensitive to slight misalignment.

We also investigated the effect of case selection on the performance of a CAD scheme, since the choice of clinical cases used to test the scheme can affect the test results. In this study, we deliberately modified the components of our database to study the effects of this modification on measured performance. Using our computerized scheme for the automated detection of breast masses from mammograms, we found that the sensitivity of the scheme ranged between 26% to 100% (at a false positive rate of 1.0 per image), depending on the cases used to test the scheme. Even a 20% change in the cases comprising the database reduced the measured sensitivity by 15-25% (ref. 33). Because of the strong dependence of measured performance on the testing database, it is difficult to estimate reliably the accuracy of a CAD scheme. Moreover, it is questionable to compare different CAD schemes when different cases are used for testing. Sharing databases, creating a common database, or using a quantitative measure to characterize databases are possible solutions to this problem. However, none of these solutions exists or is practiced at present. Therefore, as a short-term solution, we recommend that the method used for selecting cases and histograms of relevant image features be reported whenever performance data are presented.

The importance of our findings is that a nonlinear bilateral subtraction technique can detect mammographic masses at a high level of sensitivity that are again comparable to levels obtained by radiologists.

(3)   Computed Detection of Lesions Missed by Mammography

Over the past 6 years, we have been collecting cases in which a lesion was missed in a mammogram. To date, 69 cases with a lesion that went undetected by a radiologists were analyzed by the two detection schemes -- clustered microcalcifications and masses (ref. 37). In all cases the lesions were rated retrospectively as being subtle to extremely subtle by an experienced radiologist. The computer schemes correctly identified approximately 50% of the missed lesions -- 54% of the malignant lesions and 45% of the

10

benign lesions. The false positive rate was 1.3 per image. This result shows that our computer detection schemes are capable of identifying cancers that are overlooked by radiologists.

(4) Classification Schemes

We have developed a method for differentiating malignant from benign clustered microcalcifications in which image features are both extracted and analyzed by a computer. One hundred mammograms obtained from 53 patients who had biopsies for suspicious clustered microcalcifications were used. Our technique used 8 computer-extracted features of clustered microcalcifications that were merged by an artificial neural network. Features were based on the size and shape of clusters and on the size, shape, contrast, and uniformity of individual microcalcifications comprising a cluster. Human input was limited to initial identification of the microcalcifications. Our method correctly classified 100% of patients with breast cancer and 69% of patients with biopsy-proven benign conditions. ROC analysis showed that our method performed significantly ($p=0.03$) higher than 5 radiologists who reviewed the mammograms retrospectively. This result indicated that quantitative features extracted by a computer can be analyzed by a computer to distinguish malignant from benign clustered microcalcifications, and that our technique can potentially help radiologists to reduce the number of "false-positive" biopsies.

2. BODY

2.1. Experimental Methods, Assumptions and Procedures

The overall plan of this demonstration project involves four major steps, namely, (1) further development of advanced CAD schemes, (2) development of prototype mammography CAD workstations, (3) clinical evaluation of prototype workstations, and (4) analysis of outcomes from clinical evaluations.

11

The primary goal of this study is to demonstrate that approximately 23 additional breast cancers will be detected by the use of prototype mammography CAD workstations for approximately 45,000 screenees who are expected to enter a three-year clinical evaluation at two hospitals. The potential of detecting 23 additional breast cancers was estimated from an average breast cancer incidence rate of five per 1,000 screenees, a current average miss rate of 20%, and a level of CAD performance that detects 50% of currently missed cancer lesions.

Advanced CAD schemes will be developed for detection of clustered microcalcifications and masses as well as characterization of detected lesions by integrating a number of new methods into the existing programs and optimizing a number of parameters for achieving high performance levels above the current ones. Two kinds of prototype mammography CAD workstations will be developed. The first prototype unit is based on the existing intelligent workstation at the University of Chicago, which will incorporate the most advanced CAD software and will be used for clinical evaluation on approximately 30 screenees per day at the University of Chicago. The second type is the prototype commercial units which will be developed by R2 Technology, Inc., and will be used for clinical evaluation on approximately 30 screenees per day at LaGrange Memorial Hospital.

The impact of the computer output from the prototype workstation will be evaluated by examining if and when the radiologist changes his/her initial diagnosis. The computer output will be presented to the radiologist only after the radiologist has entered his/her initial findings into the computer as to the normal and abnormal lesion(s). A particularly important datum in this demonstration project is the measurement of the number of breast cancer cases on which the radiologist did not initially indicate the breast cancer lesion but did make a final correct diagnosis by using the computer output as a "second opinion."

In this demonstration project, we will not direct effort toward the development of major new methods and techniques on mammographic CAD schemes. Instead, we plan to incorporate several useful methods and techniques, which are recently developed, into the CAD software package for implementation in the prototype intelligent mammography workstation. It is important to note that considerable research effort would be required to optimize many parameters associated with new CAD methods and the existing CAD algorithms in order to integrate all of the components into a single package that functions successfully.

In the first phase of this project, we plan to develop advanced CAD schemes for detection of clustered microcalcifications and masses, and then to incorporate them into the prototype intelligent mammography workstation for clinical evaluation at the University of Chicago. However, as the performance of advanced CAD schemes in our laboratory improves through continued efforts on the optimization process, the CAD software package in the workstation will be upgraded as needed. In the second phase of this project, we plan to incorporate additional CAD schemes to characterize detected lesions as benign or malignant.

(1)   Automated scheme for detection of clustered microcalcifications

We plan to investigate and incorporate three new approaches to improve the performance of automated detection of clustered microcalcifications. They are (1) local edge-gradient analysis techniques for reduction of false-positives, (2) shift-invariant neural networks for removal of false-positives, and (3) wavelet transform techniques for improvement in the sensitivity in detecting clustered microcalcifications, as described below. Many parameters associated with these approaches will be selected carefully to optimize the overall performance in detecting clustered microcalcifications. It is important to note that previous studies on these methods were based on mammograms digitized using a drum scanner. In this project, we plan to

13

determine all of the new parameters with mammograms digitized using a laser scanner that is integrated into the prototype intelligent mammography workstation.

(2)   Automated scheme for detection of masses

We plan to investigate and incorporate three new approaches to improve the performance of automated detection of mass lesions: (1) Hough spectrum analysis for the detection of spiculated lesions and architectural distortions; (2) gradient and circularity analysis for the detection of very small early cancers; and (3) artificial neural networks for the merging of various features of suspect lesions, identified either by the bilateral subtraction method or by the two new single image methods, in order to reduce the number of false-positive detections.

(3)   Automated scheme for characterization of detected lesions

In the second phase of development of advanced CAD schemes, we plan to investigate and incorporate two automated schemes for distinguishing between benign and malignant lesions both for detected clustered microcalcifications and masses. The likelihood of malignancy on each detected suspicious lesions will be calculated from our schemes and will be displayed together with the location(s) of detected lesion(s) on the prototype mammography CAD workstation at the University of Chicago. We plan to investigate whether the calculated likelihood of malignancy added to the CAD computer output may improve the diagnosis of breast cancer by reducing the false-positives and false-negatives.

(4)   Development of prototype mammography CAD workstations

We plan to develop two kinds of prototype mammography CAD workstations for clinical evaluation. One is based on the existing intelligent mammography workstation at the University of Chicago, which will be modified by incorporating advanced CAD

14

software and by improving some aspects of the hardware configuration. This first prototype system will be used for clinical evaluation at the University of Chicago. The second type of prototype system will be developed by R2 Technology, Inc., as a potential commercial unit, and will be placed for clinical evaluation at LaGrange Memorial Hospital. Although the basic principles employed in the two kinds of prototype workstations are similar due to licensing of the University of Chicago technologies to R2 Technology, Inc., these two systems are not identical. Therefore, we plan to investigate the levels of performance of each prototype workstation.

(5)   Clinical evaluation of prototype mammography CAD workstations

Multi-institutional clinical evaluation of mammography CAD workstations will be carried out at two clinical sites: the Mammography Section of the Department of Radiology, the University of Chicago and LaGrange Memorial Hospital in LaGrange, Illinois. The number of screenees per day who will enter this clinical evaluation at each of the two hospitals is approximately 30. The total number of screenees per day will be 60. We have already obtained an approval from the Institutional Review Board (IRB) for clinical evaluation of the prototype intelligent mammography workstation at the University of Chicago and LaGrange Memorial Hospital.

To examine the impact of mammographic CAD on clinical outcomes, we plan to obtain data from mammography audits without and with the prototype CAD workstations. For the first six months of this project, the CAD workstation will not be used and we will collect results of mammography audits. For the next year, the first clinical evaluation with the CAD workstation will be carried out. Then, a second mammography audit will be conducted for the subsequent six-months period without the CAD workstation. We believe that this second segment will be useful to obtain additional baseline data and also to examine the potential variation in the baseline data without the CAD workstation being used clinically. For the final two-year period, the

15

second clinical evaluation of the CAD workstations will be carried out. We will audit the total of three-year periods when the CAD workstations were used and compare those results to the audit of the two six-month audits. This will allow us to study the effects of CAD by comparing parameters such as sensitivity, call-back rates, positive predictive value, etc.

For daily clinical evaluation of the CAD workstations, all screening mammograms will be digitized by a research technologist at each of the two sites and the computed results from the CAD schemes will be stored. When the radiologist reads the original film mammogram, he/she enters his/her findings on normal or abnormal lesion(s) into the CAD workstation using a light pen and soft copy of the mammograms on CRT monitors. Then, the computer output will be indicated on the monitor. The radiologist will then have an opportunity to modify his/her opinion using the light pen. If the radiologist changes his/her initial diagnosis due to the computer output, then the radiologist will enter the final result into the computer. With this procedure, we will be able to determine the number of breast cancer cases on which the radiologist may miss the lesion initially but may correct his/her findings using the CAD output.

(6)   Analysis of outcomes from clinical evaluation of prototype mammography CAD
       workstations

The effect of mammography CAD workstations on clinical outcomes in the detection of breast cancer will be analyzed both prospectively on a daily basis using the workstation and on a semi-annual basis using the results of mammography audits. Radiologists' performance will be evaluated as a group and also as individuals in order to examine the inter- and intra-observer variability. Since each of the two clinical sites has already established its own mammography audit system, data for "truth" in terms of normal/abnormal (breast cancer) cases will be obtained from each site's mammography audit system for analysis of outcomes in this demonstration project.

16

2.2 Results and Discussion

(1) Development of automated detection scheme for clustered microcalcifications

We have been developing techniques for optimizing our rule-based scheme. Previously, we investigated the use of a genetic algorithm for selecting the optimum set of thresholds for our detection scheme. The genetic algorithm used a cost function that combined the false-positive rate and the true-positive rate to produce a single value. This required us to arbitrarily assign weightings to true-positive versus false-positive rates, which is a very difficult task. Using this approach, a single set of thresholds corresponding to one set of true-positive and false-positive rates (a single operating point on an FROC curve) was obtained. However, if the weightings assigned to the true- and false- positive rates were not the best choice, then the solution of the genetic algorithm would not be optimal. Because only a single operating point was obtained, it is difficult to assess the appropriateness of this solution. We are now investigating the use of a multi-objective genetic algorithm, which produces an optimum FROC curve, not just a single operating point. Using the previous genetic algorithm approach an optimum solution corresponding to 87% sensitivity at 1.0 false positives per image was obtained. Using the multi-objective approach, the same operating point was obtained, in addition to another of other points. For a lower false-positive rate, say 0.2 per image, a sensitivity of 83% can be obtained. A higher sensitivity, say 95%, can be obtained at 2 false positives per image. We believe that the multi-objective genetic algorithm is the best approach for optimizing our scheme.

We have investigated the use of a multi-objective genetic algorithm (MOGA) to optimize our rule-based scheme. The MOGA is a search method to find the optimal set of sensitivity and specificity pairs by efficiently searching through the set of all possible solutions. The initial study of the MOGA was done on our standard dataset for development of techniques. We are now in the process of optimizing our detection

17

scheme using the MOGA on a set of cases from our clinical database, augmented with an additional 50 cancer cases selected from our film library.

(2)   Development of automated detection scheme for masses

We have incorporated three techniques to improve the overall performance of our CAD schemes for detection of masses. Three techniques include Hough spectrum analysis, gradient and circularity analysis, and artificial neural networks. We attempted to achieve the high overall performance by optimal selection of many parameters involved in this scheme and also to examine various classifiers to distinguish between lesions and false positives. In our CAD scheme, many features are extracted from potential lesion sites and merged into a single decision variable using a classifier. Numerous features can be extracted from potential lesion sites making it difficult to optimally choose representative features to be used as inputs to a classifier. We have undertaken the problem of feature selection for two different classifiers using a dataset consisting of features extracted from lesions and false-positive detections. We have applied traditional feature selection techniques such as single feature selectors and stepwise selectors. In addition, we have applied genetic algorithms to this search task. A genetic algorithm is an optimization technique loosely based on natural selection. Multiple solutions to a problem are randomly generated and their "fitness" is evaluated. Solutions with better fitness values are more likely to survive to subsequent generations, while solutions with a poor fitness value will "die out." This "survival of the fittest" strategy usually results in a rapid convergence to the optimal solution. By employing genetic algorithms, we have improved the A_Z of our mass CAD scheme from 0.96 to 0.98 using artificial neural networks. With linear discriminants, the A_Z improved from 0.93 to 0.95. The results from the linear discriminant analysis show that the genetic algorithm feature selection method is as good as, if not better than the stepwise method. Similar results were obtained for the artificial neural network classifiers but the results were not

as strong. As with all studies employing neural networks, it is possible that there is over-fitting of the data. We attempted to minimize this effect by simplifying the structure of our networks and by employing cross-validation or leave-one-out tests.

A new development, which is now being implemented into the mass detection scheme, is a new region growing algorithm. We have developed two novel lesion segmentation techniques -- one based on a single feature called the radial gradient index (similar feature to that described above) and one based on a simple probabilistic model to segment mass lesions from surrounding background. In both methods a series of image partitions is created using gray-level information as well as prior knowledge of the shape of typical mass lesions. With the former method the partition that maximizes the radial gradient index is selected. In the latter method, probability distributions for gray-levels inside and outside the partitions are estimated, and subsequently used to determine the probability that the image occurred for each given partition. The partition that maximizes this probability is selected as the final lesion partition (contour). We tested these methods against our previous region-growing algorithm using a database of biopsy-proven, malignant lesions and found that the new lesion segmentation algorithms more closely match radiologists' outlines of these lesions. At an overlap threshold of 0.30, gray level region growing correctly delineates 62% of the lesions in our database while the radial gradient index algorithm and the probabilistic segmentation algorithm correctly segment 92% and 96% of the lesions, respectively. With these new segmentation results we hope to find and extract new features that will help differential between actual lesions and false-positive detections, thus improving the overall performance of computerized mass detection.

Our computerized detection method for masses initially identifies various suspect lesion sites. Features from these sites are then extracted automatically and merged by a classifier in order to reduce the number of false-positive detections. Different subsets of features will, in general, result in different classification performances. We investigated

the effect of having a limited datasets on feature selection. We showed that, with limited datasets and/or a large number of features from which to choose, bias is introduced if the classifier parameters are determined using the same data that were employed to select the "optimal" set of features.

We have investigated the use of a Bayesian neural network in the merging of computer-extracted features of actual lesions and false-positive detections. We found that with a limited dataset, use of the Bayesian network reduces the potential for overtraining typically encountered in conventional neural networks.

We have previously investigated the use of a radial gradient index (RGI) to aid in the segmentation of mass lesions from parenchymal background in digitized mammograms. In this work, we develop a non-linear filtering algorithm based on the RGI that creates RGI feature images from digital mammograms, which can be subsequently thresholded to distinguish between mass lesions and normal regions. This initial stage of mass detection is focussed on improving sensitivity leaving later feature analysis and classification stages for reducing false-positive detections. Using just RGI filtering, we achieved a sensitivity of 93% with 16 false detections per image on a database of 60 patients (112 images). After feature analysis and classification on the suspect regions, the by-patient sensitivity of 77% at 2 false positives per image was obtained.

(3)   Development of automated scheme for characterization of clustered microcalcifications

We have developed an automated scheme for the classification of clustered microcalcifications as malignant or benign. We have shown in previous studies that this computer scheme to be more accurate than radiologists in differentiating between malignant and benign microcalcifications. We have also shown in an observer study that this computer aid can help radiologists improve their diagnostic accuracy and improve their biopsy recommendations. In this observer study, ten radiologists read the

20

mammograms from 104 patients with and without our computer aid, and they reported their confidence that a microcalcification cluster represented a malignancy and also reported their recommendation of biopsy or follow-up. We performed two additional analyses of the observer study data to investigate the effects of the computer classification scheme on radiologists' diagnostic performance.

In one analysis, we compared the variability with and without our computer aid in the radiologists' interpretation of malignant and benign microcalcifications and in their recommendations for biopsy or follow-up. First, when the computer aid was used, variation in the radiologists' diagnostic accuracy as measured by the standard deviation of the area under the ROC curves (A_Z) was reduced 47%. This reduction in variability is in addition to a statistically significant gain in diagnostic accuracy, as measured by (1) an increase of the average of A_Z from 0.61 to 0.75 ($p<0.0001$), (2) an increase of 6.4 biopsies per radiologist on cancer cases ($p=0.0006$), and (3) a decrease of 6.0 biopsies per radiologist on benign lesions ($p=0.003$). Second, use of the computer aid increased the agreement by all ten observers from 13% to 32% of total cases ($p = 0.0002$). The kappa statistic which is a quantitative measure of agreement, increased from 0.19 to 0.41 ($p<0.05$). Finally, use of the computer aid eliminated two thirds of substantial disagreements where biopsy and routine screening were recommended for the same patient by two radiologists ($p<0.05$). We conclude that CAD holds the potential to reduce the variability in radiologists' interpretation of mammograms in addition to its potential to improve diagnostic accuracy.

In the second analysis, we reviewed those cases that the radiologists' recommendation of biopsy or follow-up was altered by the computer aid. These consisted of 31% of the total cases. Radiologists were more likely to recommend additional biopsies when the computer estimated high values of likelihood of malignancy, and they were more likely to drop biopsy recommendations when the computer estimated low values of likelihood of malignancy. The overall probability of recommending an

additional biopsy was similar to the overall probability of dropping a biopsy recommendation (15% versus 16%). The probability of recommending an additional biopsy with a high computer-estimated likelihood of malignancy was similar for malignant cases and for benign cases (26% versus 28%). However, the probability of dropping a biopsy recommendation with a low computer-estimated likelihood of malignancy was much higher for benign cases than for malignant cases (39% versus 22%). We conclude that CAD can be used to improve radiologists' ability to differentiate between malignant and benign clustered microcalcifications and to improve radiologists' biopsy recommendations.

We have developed an automated scheme for the classification of clustered microcalcifications as malignant or benign. We have shown in previous studies that this computer scheme to be more accurate than radiologists in differentiating between malignant and benign microcalcifications. We have also shown in an observer study that this computer aid can help radiologists improve their diagnostic accuracy and improve their biopsy recommendations. In this observer study, ten radiologists read the mammograms from 104 patients with and without our computer aid, and they reported their confidence that a microcalcification cluster represented a malignancy as well as their recommendation of biopsy or follow-up.

To compare computer-aided diagnosis (CAD) with double readings by radiologists, we conducted a comparative study using data from the observer study. We derived radiologists' double-reading performance post hoc from their independent and unaided single reading data using five different objective rules of independent double readings and another rule of simulated-optimal double reading that assumed that consultations for resolving two radiologists' different independent diagnoses always produce the correct clinical recommendation. From these results and the unaided single reading and CAD reading data, we calculated sensitivity and specificity from the

observers' biopsy recommendations and obtained ROC curves from their diagnostic confidence ratings.

We found that the unaided single reading yielded 74% sensitivity and 32% specificity; whereas the CAD reading had 87% sensitivity, 42% specificity, and appeared on a higher ROC curve than the unaided single reading ($p < 0.0001$). Five methods of formulating independent double readings generated sensitivities between 59% and 89% and specificities between 50% and 13%, with their resulting operating points appearing essentially along the unaided single-reading ROC curve. The result of the simulated-optimal double reading, however, was similar to that of CAD reading, with 89% sensitivity and 50% specificity.

We conclude from this study that no real-world combinations of double reading improves diagnostic performance except for CAD reading, which approaches the simulated optimal performance.

We have developed an automated computer technique that classifies clustered microcalcifications in mammograms as malignant or benign. We have shown previously in two observer studies that this computer technique can both be more accurate than radiologists and help radiologists to be more accurate in differentiating benign from malignant clustered microcalcifications. This computer classification technique was developed on digitized screen-film mammograms. In an effort to develop this technique for full-field digital mammograms, we conducted a study of this technique on small-field digital mammograms obtained during stereotactic biopsy procedures. The goal of this work was not to analyze these small-field mammograms per se, but to analyze these mammograms that were obtained with a digital detector. The rationale is that we expect the findings from this analysis of the small-field digital mammograms to apply, in principle, to full-field digital mammograms as well. In this study, we analyzed 79 lesions, of which 33 were malignant and 46 were benign. Each of these cases typically consisted of more than one image and, therefore, we analyzed a total of 176 images, of

which 56 were of the malignant lesions and 120 were of the benign lesions. We applied the same computer technique developed based on digitized screen-film mammograms using the same computer-extracted image features. The computer technique achieved an A_z value of 0.84 for the 176 images and 0.90 for the 79 lesions. In comparison, radiologists who evaluated these lesions prior to biopsy achieved an A_z value of 0.76 for the 79 lesions. Therefore, our computer technique outperformed the radiologists in classifying these breast lesions as malignant or benign. We concluded from this study that our computer technique can potentially classify clustered microcalcifications accurately as malignant or benign in mammograms acquired with digital detectors.

(4)   Development of automated scheme for characterization of masses

The automated classification of masses begins by segmenting the lesions using a grey-level region growing applied to a 512x512 ROI (region of interest) centered on the lesion after background-trend correction (using a second order polynomial) and histogram equalization. The grey-level threshold value is determined from a "transition point." The transition point is the grey level for which there is a discontinuous decrease in the circularity and a corresponding discontinuous increase in size of the grown lesion (ref. 39).

From the segmented lesion, four features related to the degree of spiculation, margin sharpness, density of each mass, and the texture within the mass are extracted automatically from the neighborhoods of mass regions. The techniques for extracting these four features are described in ref.(39). Because of its strong ability to differentiate benign from malignant masses, degree of spiculation is first used in a rule-based technique (i.e., a threshold is applied to the degree of spiculation measure). Those masses that have a spiculation measure lower than a threshold value are then subjected to the ANN, where the remaining features are used as input. The architecture of the ANN is

24

three input units, two hidden units, and one output unit. The spiculation measure and the output of the ANN are used to determine the likelihood of malignancy.

Using a database of 95 mammograms containing masses from 65 patients (all but one having been biopsied for the suspicion of breast cancer), the performance of the mass classification technique was measured and compared to the results of interpretations by radiologists reading the same cases. Using ROC analysis, the computer classification scheme yielded an A_z value of 0.94, similar to that of an experienced mammographer (A_z=0.91) and statistically significantly higher than the average performance of the radiologists with less mammographic experience (A_z=0.80). With the database we used, the computer scheme achieved, at 100% sensitivity, a positive predictive value of 83%, which was 12% higher than that of the experienced mammographer and 21% higher than that of the average performance of the less experienced mammographers at a p-value of less than 0.001.

The robustness of the computerized scheme to case-variation was evaluated on an independent database consisting of 110 new cases (50 malignant and 60 benign). Mammograms from the independent database were digitized twice using two different laser scanners (Konica LD 4500 and Lumiscan 100) in order to evaluate the robustness of the scheme to the variation in digitization techniques. Using ROC analysis, the classification scheme achieved A_z values of 0.82 (Konica) and 0.81 (Lumiscan) on the independent database. Results from statistical analyses showed that the differences in the performance due to the case-variation between the training and independent databases and the variation in film digitization techniques were not statistically significant (p=0.14, 0.10 and 0.76). The independent evaluation of the computerized scheme for the classification of benign and malignant masses showed that the computerized classification scheme is robust to the variations in case-difficulty and digitization techniques.

In our computerized classification method for estimating the likelihood of malignancy of mammographic masses, we investigated two different classifiers -- an artificial neural network (ANN) and a hybrid system (one stop rule-based followed by an artificial neural network). In order to understand the difference between the two classifiers, we investigated their learning and decision-making processes by studying the relationships between the input features and the outputs. A correlation study showed that the outputs from the ANN-alone method correlated strongly with one of the input features (spiculation) ($r = 0.91$), whereas the correlation coefficients for the other features ranged from 0.19 to 0.40. This strong correlation between the ANN output and spiculation measure indicates that the learning and decision-making process of the ANN-alone method were dominated by the spiculation measure. We found that with a limited database, it is detrimental for an ANN to learn the significance of other features in the presence of a dominant feature. Our hybrid system, which initially applied a rule concerning the value of the spiculation measure prior to employing an ANN, prevents over-learning from the dominant features and performed better than the ANN-alone method in merging the computer-extracted features into a correct diagnosis regarding the malignancy of the masses.

The effectiveness of the computerized classification scheme as an aid to radiologists in the task of differentiating between benign and malignant masses was evaluated in a preliminary observer study. The preliminary observer study was conducted including 20 selected cases and 128 radiologists. For each case, the observer viewed the CC, MLO and special views (e.g., magnified or spot compression view) of the mass lesion on a monitor, along with a minified image of all four standard views. The observers were asked to give their confidences regarding the likelihood of malignancy for each case, first without and then with the computer output of an estimated likelihood of malignancy. As many as 6 training cases were shown to the observers before the actual study. The 20 cases were randomized differently for each observer. The average

26

performance of the radiologists in terms of Az value was 0.89 and 0.94 without and with the computer aid, respectively. Results from the paired t-test showed that the difference in Az was statistically significant (p-value < 0.0001). The preliminary results from an observer study showed that a significant improvement in the performance of radiologists was achieved in the classification of benign and malignant masses when computer aid was used.

We evaluated our computerized classification method, which was initially developed on digitized screen/film mammograms, on a large database of digital mammograms. We collected 110 prospective cases (212 images) from a LORAD stereotactic imaging system that had initially been obtained for needle localization or core biopsy of a suspect mass lesion. The database consisted of 44 malignant cases and 66 benign cases. The computer classification method includes the automated segmentation of the mass lesions from the breast parenchyma, the automated extraction of lesion features, and the automated classification of the suspect lesion into an estimate of the likelihood of malignancy. A Bayesian neural network (BANN) was used to merge the four features of spiculation, margin sharpness, average gray level, and texture. The BANN uses regularization to prevent overtraining of the network. The untrained computer method from the screen/film database yielded an Az of 0.72 on the digital mammography database. After retraining of the BANN, the Az increased to 0.91, similar to that obtained from the radiologists' suspicion ratings of the lesions (0.92). Further investigation of the features showed that the spiculation feature performed better on the screen/film database, whereas the texture feature performed better on the digital mammography database. Due to differences in the physical characteristics of the two image acquisition systems, features values and the merging of these by classifiers needs to be carefully optimized.

To evaluate the effectiveness of a computerized classification method as an aid to radiologists reviewing clinical mammograms for which the diagnoses were unknown to

both the radiologists and the computer. Six mammographers and 6 community radiologists participated in an observer study. These 12 radiologists interpreted, without and with the computer aid, 110 cases that were unknown to both the 12 radiologist observers and the trained computer classification scheme. The radiologists, performances in differentiating between benign and malignant masses without and with the computer aid were evaluated using ROC analysis. Two-tailed p-values were calculated for Student's t-test to indicate the statistical significance of the differences in performances with and without the computer aid.

When the computer aid was used, the average performance of the 12 radiologists improved, as indicated by an increase in Az from 0.93 to 0.96 (p-value=0.0002), by an increase in 0.90A'z from 0.56 to 0.72 (p-value=0.0002), and by an increase in sensitivity from 94% to 98% (p-value =0.022). No statistically significant difference in specificity was found between readings with and without computer aid (=-0.014; p-value=0.46; 95% CI =(-0.054, 0.026)). When we analyzed results from the mammographers and community radiologists as separate groups, a larger improvement was demonstrated for the community radiologists. Computer-aided diagnosis can potentially help radiologists improve their diagnostic accuracy in the task of differentiating between benign and malignant masses seen on mammograms.

(5)   Development of prototype CAD workstation

Our intelligent workstation consists of an IBM RISC 6000 Powerstation 590, a Konica LD4500 film digitizer, an Alphatronix Inspire magneto-optical jukebox, 2 Imlogix 1000 CRT monitors and a Seikosha VP4500 thermal printer. The system has been used in the clinical reading area of the Department of Radiology since November 8, 1994.

Each day all screening mammograms (4-views per case) were digitized. As the films are being digitized, using a 100 micron pixel size and 1024 grey levels, the

microcalcification detection program is run on-line in parallel. The mass detection program is run off-line overnight, since the films are not reviewed until the next day. After all four films have been analyzed, the results of the microcalcification detection program are displayed in a single 1024x1280 image as a collage of four 512x620 images with arrow(s) displayed on the image as annotation indicating the computer results. The results were then recorded on thermal paper, upon which the radiologists can make notes and comments. The results of the mass detection program were printed using the same format the next morning. A full case, four films, can be processed in less than 5 minutes.

Recently, we have made a major modification in the existing workstation by incorporating a touch-screen CRT monitor to display the results of the computer analyses to the radiologist. This replaced the thermal paper copy and will facilitate recording of radiologists' findings. The touch-screen system is used for recording the location of lesions that the radiologist believe are malignant. A digital copy of the four views is displayed on a monitor with no computer results. The radiologist, after reading the original film mammograms, touches the screen of the CRT monitor to indicate region(s) in the images that may contain cancer. If the radiologist considers no cancer lesion to be present in the image, he/she also enters this initial normal finding to the workstation using the touch screen, using a button displayed on the CRT monitor outside the breast region. Once this is done, the computer results are displayed on the CRT monitor and the radiologist, after reviewing the computer results together with the original films, again uses the touch screen to indicate suspicious region(s) in the images on the monitor, if the location of the malignant region (s) found with the computer output is different from the initial location, or if the initial finding is normal.

(6)   Clinical evaluation of CAD workstations

As of December 2001, over 25,000 cases have been analyzed by using our CAD workstation at the University of Chicago. We are analyzing the sensitivity and false-

positive rate of the workstation for the first three years (12,670 cases). With follow-up of up to five years for some patients, 79 women have developed breast cancer in our study cohort. Of the 79 cancers, 61 were initially detected on a screening mammogram. The remaining cases were initially detected either on a diagnostic mammogram or by physical examination. Of these, 14 had true negative screening mammograms even in retrospective review, and 4 were read as negative, the cancer was visible in retrospect. Of the 65 mammographically-visible cancers, the computer identified the cancer in 44 cases (31/46 for masses and 13/19 for calcifications). For the 79 cancer patients, 42 had a negative screening mammogram that was included in our study cohort. Retrospective review of these cases showed that 19 were mammographically occult. In the 23 cases that had a subtle lesion visible in retrospect, the computer identified 12 of them. Thus, the computer was able to detect 52% of "missed" cancers approximately one year prior to diagnosis. The cases containing a missed cancer were used in an observer study to see if radiologists can detect more cancers when they use the computer aid.

The false positive rate of the computer schemes increased is currently 2.15 false masses per image and from 1.0 false clusters per image. The clustered microcalcification false positive rate decreased from approximately 1.7 to 1.0 when the screening clinic moved to a new location within the hospital. It appears that the new darkroom is cleaner than the old one and therefore there are now less film artifacts in the images.

The R2 Technology M1000 ImageChecker was installed at Grant Square Imaging in early April 1998 to support the Demonstration Project. Since that time, all mammographic interpretations performed there have been done with computer assistance. Installation of a new radiology information system at the site at approximately the same time has facilitated data collection. In addition to basic mammography audit data, the radiologists also recorded all cases in which computer assistance altered patient care, most typically resulting in a call back for computer detected finding.

The baseline data for interpretation of mammograms without computer assistance at Grant Square extends from 1-1-97 until 3-31-98. All mammographic interpretations from 1-1-97 to 12-31-97 corresponding to BIRADS categories 4 and 5 have been tracked to this point. Results for this year will be finalized after physicians offices have been contacted a third time about several cases.

In a positive development that should add greatly to the number of examinations included in the study, LaGrange Memorial Hospital has decided to purchase an M1000. The baseline period for interpretation of mammograms without computer assistance at LMH will be 1-1-97 to approximately 12-31-98. The audit for radiologists' performance for 1-1-97 to 12-31-97 is essentially complete. Highlights include: volume of approximately 7500 cases; a 4.1 per thousand cancer detection rate; a 73% minimal cancer detection rate (Tis, T1a and T1b lesions); a 7% call back rate. The protocol for procedure interpretation with computer assistance will be the same at Grant Square and LMH.

The R2 Technology ImageChecker was installed at Grant Square Imaging in early April 1998 to support the demonstration project. A number of conventional mammography practice parameters were monitored between January 1, 1997 and March 31, 2000 in order to assess the effect and utility of the ImageChecker. Additionally, commencing on August 1, 1998, radiologists were asked to record cases in which patients were recalled to the department for additional imaging as the result of CAD prompts.

The volume of mammography at the clinic averaged between 10 and 12 cases per day throughout the course of the study. A single radiologist staffed the site and, in addition to mammograms, also interpreted plain film studies, as well as computed tomography, nuclear medicine and ultrasound examinations with the total case volume being approximately 40 studies per day. The volume and mix of the examinations was such that radiologists at the clinic were about "half as busy" as at other sites staffed by

31

the radiology professional group. Four radiologists read more than 95% of the mammograms at the clinic during the 39 month study period.

It was the policy of the clinic that no added charges be assessed for additional mammography views. The technologists therefore routinely reviewed the cases before patient discharge and brought potential abnormalities to the radiologist's attention who then decided whether additional views might be needed before the examination was considered complete. This practice style did not change after installation of the ImageChecker, with technologist review coming before the films were digitized for CAD analysis. The approach described tended to reduce significantly the percentage of patients "called back" to the department on a separate date for additional views. It also reduced the potential for observational oversights due to a single reading by one radiologist.

Between August 1, 1998 and March 31, 2000, 51 of 5359 (1.0%) patients were called back to the department based on prompts by the ImageChecker. An independent assessment of the average recall rate for the three radiologists who interpreted examinations throughout the entire study period showed an increase from 3.1% between January 1, 1997 and March 31, 1998 (before installation) to 4.2% between the August 1, 1998 and March 31, 2000 (p=0.037; Pearson Chi-square test with Yates correction). As described above, the practice style at the clinic tended to reduce recall rates appreciably.

The 51 additional patient recalls resulted in 6 biopsies, one of which yielded a malignant diagnosis. This represented an increase of 5.2% (1/19) in the yield of impalpable cancers at the clinic during the monitored period

No statistically significant difference in the detection rate for impalpable breast cancer was noted before and after ImageChecker installation. Between January 1, 1997 and March 31, 1998 the detection rate was 4.5/1000 (13/2866). Between April 1, 1998 and March 31, 2000 the detection rate was 3.8/1000 (24/6345), (p=0.73). The positive predictive values for biopsy before and after installation were 38% and 30% respectively

32

(p=0.55). In cases for which T staging was available, the minimal cancer detection rate (percentage Tis, T1a and T1b lesions) was 50% (4/8) before installation and 75% (9/12) after installation (p=0.36; 2-tailed Fisher exact test).

The practice setting in this study is one that could be reasonably expected to minimize the benefits of computer-aided detection. The radiologists reviewed a relatively small number of cases every day and did so in a rather relaxed environment. Additionally, the technologists prescreened examinations for potential abnormalities in order to reduce recall rates. In so doing, they often served effectively as second readers, an approach taken at some institutions to improve the sensitivity of screening mammography. It is therefore of interest that the ImageChecker alerted radiologists to abnormalities, overlooked on original film review, but warranting patient recall in nearly 1.0% of the cases. Further, biopsy recommendations resulted in more than 10% of the recalled patients (6/51). These results suggest that even in a favorable practice environment, radiologists do not perceive all findings that, retrospectively, can be considered worthy of concern.

The question does arise whether CAD information is as beneficial in a low-volume reading situation. In particular, one could postulate that CAD prompts might make the unhurried radiologist hyperattentive to borderline findings that ultimately prove to be benign. It has been shown that such borderline findings; at least when identified by experts, almost never are the result of malignant lesions. The percentage of patients additionally recalled due to CAD prompts in this study was approximately 1.0%, a value similar to that reported by Freer and Ulissey. At the same time, the incremental increase in yield for breast cancer detection was much less in this study, i.e., 5%, than the nearly 20% reported by Freer and Ulissey in a study carried out in a much busier mammography center. Given the relatively small number of cancer cases involved, however, the difference is not statistically significant (p=0.27; 2-tailed Fisher exact test). Radiologists

using CAD information in any setting should nevertheless be aware of the very low yield in the recall of borderline findings.

There are obvious difficulties in arriving at statistically meaningful observations in low-case-volume practice. We were fortunate to have a very stable practice environment for the 39 month study period during which slightly more than 9000 examinations were performed. Given the very low incidence of breast cancer, however, our results fail to prove or disprove the benefit of CAD in terms of the most reliable indicator, i.e. absolute detection rates at screening. Such a benefit, if any, may be difficult to prove satisfactorily at centers that perform a low volume of the examination.

In summary, the utilization of the R2 ImageChecker at a low volume mammography center resulted in an apparent small (5.0%) increase in the detection of impalpable breast cancer while an additional 1.0% of patients were recalled for additional imaging as a result of CAD prompts. The benefits of CAD at clinics doing a low volume of mammography may be less than those obtained at high-volume centers.

2.3 Address to the Statement of Work

The original statement of work in the revised application is listed below. All of the proposed work has been completed.

Task 1: Months 1-36: Development of advanced CAD schemes for detection of breast lesions

> Task 1A: Months 1-36: Development of automated detection scheme for clustered microcalcifications

> Task 1B: Months 1-36: Development of automated detection scheme for masses

> Task 1C: Months 1-36: Development of automated schemes for characterization of detected breast lesions

Task 2: Months 1-48: Development of prototype CAD workstations

> Task 2A: Months 1-6: Development of U of C's workstation

34

Task 2B: Months 1-6: Development of R2 Technology's workstation

Task 2C: Months 7-48: Refinement and upgrade of workstations at the University of Chicago

Task 2D: Months 7-48: Maintenance and upgrade of workstations by R2 Technology, Inc.

Task 3:   Months 1-48: Clinical evaluation of CAD workstations at two hospitals

Task 3A: Months 1-6 and 19-24: Mammography audit without CAD workstations

Task 3B: Months 7-19 and 25-48: Mammography audit and evaluation with CAD workstation

Task 4:   Months 1-48: Analysis of outcomes from clinical evaluation of prototype workstations

Task 4A: Months 7-18: Collection of cases for observer studies

Task 4B: Months 19-48: Analysis inter- and intra-observer variability

Task 4C: Months 19-48: Analysis of the variability with prototype mammography CAD workstations

3.   KEY RESEARCH ACCOMPLISHMENTS

(1) We developed an advanced automated detection scheme for clustered microcalcifications based on the use of a multi-objective genetic algorithm (MOGA).

(2) We developed an advanced automated detection scheme for masses by use of a Bayesian neural network and a radial gradient index.

(3) We demonstrated that an advanced scheme for characterization of clustered microcalcifications as malignant or benign can help radiologists improve their diagnostic accuracy and also their biopsy recommendations.

35

(4) We demonstrated that CAD can potentially help radiologists improve their diagnostic accuracy in the task of differentiating between benign and malignant masses on mammograms.

(5) We developed a prototype CAD workstation for clinical evaluation, and demonstrated that the computer was able to detect 52% of missed cancers approximately one year prior to diagnosis.

(6) We evaluated the effect of the R2 ImageChecker computer-aided detection system in a low-volume mammography clinic, and demonstrated an increase in the early detection rate of impalpable breast cancer by 5.2% over a 24-month period.

4.   REPORTABLE OUTCOMES

4.1 82 papers were published as listed in 6. Bibliography of all publications.

4.2 14 patent applications have been filed as listed below:

(1) Ishida T, Katsuragawa S, Doi K: Iterative image warping technique for temporal subtraction of chest radiographs. U.S. Pat. Serial No. 09/053,798 (allowed). **(UCHI#707)**

(2) Doi K, Nakamura K: Computerized analysis of the likelihood of malignancy in solitary pulmonary nodules on chest radiographs. U.S. Pat. Serial No. 09/830,574 pending. **(UCHI#715)**

(3) Xu XW, Doi K, MacMahon H: Computerized detection of lung nodules using energy-subtracted soft-tissue and conventional chest images. U.S. Pat. No. 6,240,201. **(UCHI#716)**

(4) Doi K, Li Q, Katsuragawa S, Ishida T: Method, system and computer readable medium for computerized processing of chest radiographic images. U.S. Pat. Serial No. 09/830,562 pending. **(UCHI#714)**

(5) Li Q, Katsuragawa S, Doi K: Method, system, and computer readable medium for computerized processing of contralateral and temporal subtraction images using elastic matching. U.S. Patent Serial No. 09/692,218 pending. **(UCHI#802)**

(6) Doi K, Arimura H, Katsuragawa S, Morishita J: Computerized method for identification of chest radiographs using image mapping and template matching techniques. U. S. Patent Serial No. 09/816,217 pending. **(UCHI#823)**

(7) Li Q, Katsuragawa S, Doi K: Process, system and computer readable medium for pulmonary nodule detection using multiple-templates matching. U.S. Patent Serial No. 09/716,335 (allowed). **(UCHI#861)**

(8) Doi K, Aoyama M: Automated computerized scheme for classification of lung nodules on chest images. U.S. Patent Serial No. 09/818,831 pending. **(UCHI#885)**

(9) Doi K, Ishida T, Katsuragawa S: Novel subtraction CT technique for computerized detection of small lung nodules. U.S. Patent Serial No. 09/990,311 pending. **(UCHI#945)**

(10) Doi K, Aoyama M, Q. Li: Computerized method for determination of the likelihood of malignancy for pulmonary nodules on low-dose CT. U.S. Patent Serial No. 09/990,310 pending. **(UCHI#946)**

(11) Li Q, Doi K: Computerized scheme for distinguishing between benign and malignant nodules on chest CT by use of similar images. U.S. Patent Serial No. 09/990,377 pending. **(UCHI#947)**

(12) Suzuki K, Doi K: Massive training artificial neural network (MTANN) for detection of abnormalities in medical images. U.S. Patent Serial No. 10/120,420 pending. **(UCHI#990)**

(13) Doi K. Uchiyama Y, Katsuragawa S: Quantitative computer-aided analysis of diffuse lung diseases in high-resolution CT. U.S. Patent pending. **(UCHI#1027)**

(14) Li Q, Doi K: Selective enhancement filters for computerized detection of pulmonary nodules in three-dimensional thoracic CT scans. U.S. Patent pending. **(UCHI#1028)**

5. CONCLUSIONS

We have made significant progress in the development of various CAD schemes for detection and characterization of breast lesions. Evaluation of our CAD workstation and collection of mammographic audit data have been completed. We believe that this project has produced a useful result concerning the impact of CAD schemes in the additional detection of breast cancer.

6. BIBLIOGRAPHY OF ALL PUBLICATIONS

(1)   Bick U, Giger ML, Schmidt RA, Nishikawa RM, Wolverton DE, Doi K: Computer-aided breast cancer detection in screening mammography. In: Digital Mammography '96 (eds. K. Doi, M. L. Giger, R. M. Nishikawa, R. A. Schmidt). Amsterdam (Elsevier Science (Excerpta Medica International Congress series), pp. 97-103, 1996.

(2).   Bick U, Giger ML, Schmidt RA, Nishikawa RM, Doi K: Peripheral density correction of digital mammographs. RadioGraphics 16: 1403-1411, 1996.

(3)   Doi K, Giger ML, Nishikawa RM, Schmidt RA: Digital Mammography '96 (Elsevier Science, Amsterdam), 1996.

(4)   Giger ML: Current issues in CAD for mammography. In: Digital Mammography '96 (eds. K. Doi, M. L. Giger, R. M. Nishikawa, R. A. Schmidt). Amsterdam (Elsevier Science (Excerpta Medica International Congress series), pp. 53-59, 1996.

(5)   Huo Z, Giger ML, Olopade OI, Wolverton DE, Zhong W, Tahoces PG, Narvid SI, Baker T, Doi K: Computer-aided diagnosis: Breast cancer risk assessment from mammographic parenchymal pattern in digital mammograms. In: Digital Mammography '96 (eds. K. Doi, M. L. Giger, R. M. Nishikawa, R. A. Schmidt). Amsterdam (Elsevier Science

(Excerpta Medica International Congress series), pp. 191-194, 1996.

(6)   Huo Z, Giger ML, Vyborny CJ, Wolverton DE, Doi K: Computer-aided

diagnosis: Automated classification of mammographic mass lesions. In:

Digital Mammography '96 (eds. K. Doi, M. L. Giger, R. M. Nishikawa, R.

A. Schmidt). Amsterdam (Elsevier Science (Excerpta Medica

International Congress series), pp. 207-211, 1996.

(7)   Huo Z, Giger ML: Integrating rules and artificial neural networks in the

classification of mass lesions in digital mammograms. Proc. World

Congress on Neural Networks '96, vol. 1, pgs. 1166-1169, 1996.

(8)   Jiang Y, Nishikawa RM, Wolverton DE, Metz CE, Giger ML, Schmidt

RA, Vyborny CJ, Doi K: Malignant and benign clustered

microcalcifications: Automated feature analysis and classification.

Radiology 198: 671-678, 1996.

(9)   Jiang Y, Metz CE, Nishikawa RM: An ROC partial area index for highly

sensitive diagnostic tests. Radiology 201: 745-750, 1996.

(10)   Jiang Y, Nishikawa RM, Metz CE, Wolverton DE, Giger ML, Schmidt

RA, Vyborny CJ, Papaioannou J, Doi K: A computer-aided diagnostic

scheme for classification of malignant and benign clustered

microcalcifications in mammograms. In: Digital Mammography '96 (eds.

K. Doi, M. L. Giger, R. M. Nishikawa, R. A. Schmidt). Amsterdam

(Elsevier Science (Excerpta Medica International Congress series), pp.

219-224 , 1996.

(11)   Kupinski M, Giger ML, Doi K: Optimization of neural network inputs

with genetic algorithms. In: Digital Mammography '96 (eds. K. Doi, M. L.

Giger, R. M. Nishikawa, R. A. Schmidt). Amsterdam (Elsevier Science

(Excerpta Medica International Congress series), pp. 401-404, 1996.

(12)   Nishikawa RM, Schmidt RA, Papaioannou J, Osnis R, Haldemann

Heusler RA, Giger ML, Wolverton DE, Comstock C, Doi K: Performance of a prototype clinical "intelligent" mammography workstation. In: <u>Digital Mammography '96</u> (eds. K. Doi, M. L. Giger, R. M. Nishikawa, R. A. Schmidt). Amsterdam (Elsevier Science (Excerpta Medica International Congress series), pp. 93-96, 1996.

(13) Nishikawa RM, Wolverton DE, Schmidt RA, Pisano ED, Hemminger BM, Moody J: A common database of mammograms for research in digital mammography. In: <u>Digital Mammography '96</u> (eds. K. Doi, M. L. Giger, R. M. Nishikawa, R. A. Schmidt). Amsterdam (Elsevier Science (Excerpta Medica International Congress series), pp. 435-438, 1996.

(14) Nishikawa RM, Papaioannou J, Collins SA: Reproducibility of an automated scheme for the detection of clustered microcalcifications on digital mammograms. <u>Proc SPIE</u> 2710: 24-29, 1996.

(15) Schmidt RA, Nishikawa RM, Osnis RB, Giger ML, Schreibman K, Doi K: Computerized detection of lesions missed by mammography. In: <u>Digital Mammography '96</u> (eds. K. Doi, M. L. Giger, R. M. Nishikawa, R. A. Schmidt). Amsterdam (Elsevier Science (Excerpta Medica International Congress series), pp. 105-110, 1996.

(16) Wei D, Nishikawa RM, Doi K: On the application of a shift invariant artificial neural network for the detection of clustered microcalcifications. In: <u>Digital Mammography '96</u> (eds. K. Doi, M. L. Giger, R. M. Nishikawa, R. A. Schmidt). Amsterdam (Elsevier Science (Excerpta Medica International Congress series), pp. 283-286, 1996.

(17) Yoshida H, Doi K, Nishikawa RM, Giger ML, Schmidt RA: An improved computer-assisted diagnostic scheme using wavelet transform for detecting clustered microcalcifications in digital mammograms. <u>Acad Radiol</u> 3: 621-627, 1996.

(18) Yoshida H, Anastasio MA, Nishikawa RM, Giger ML, Doi K: Optimally-weighted wavelet packet transform for detection of clustered microcalcifications in digital mammograms. In: <u>Digital Mammography</u> <u>'96</u> (eds. K. Doi, M. L. Giger, R. M. Nishikawa, R. A. Schmidt). Amsterdam (Elsevier Science (Excerpta Medica International Congress series), pp. 317-322, 1996.

(19) Yoshida H, Nishikawa RM, Giger ML, Doi K: Computer-aided diagnosis in mammography: Detection of clustered microcalcifications based on multiscale representation. <u>CAR '96 -- Computer Assisted Radiology</u> (eds. H. U. Lemke, K. Inamura, C. C. Jaffe, M. Vannier), pgs. 390-395, 1996.

(20) Yoshida H, Nishikawa RM, Giger ML, and Doi K: Signal/background separation by wavelet packets for detection of microcalcifications in mammograms. <u>Proc SPIE</u> 2825: 805-811, 1996.

(21) Zhang W, Doi K, Giger ML, Nishikawa RM, Schmidt RA: An improved shift-invariant artificial neural network for computerized detection of clustered microcalcifications in digital mammograms. <u>Med Phys</u> 23: 595-601, 1996.

(22) Zouras WK, Giger ML, Lu P, Wolverton DE, Vyborny CJ, Doi K: Investigation of a temporal subtraction scheme for computerized detection of breast masses. In: <u>Digital Mammography '96</u> (eds. K. Doi, M. L. Giger, R. M. Nishikawa, R. A. Schmidt). Amsterdam (Elsevier Science (Excerpta Medica International Congress series), pp. 411-415, 1996.

(23) Giger ML, Nishikawa RM, Kupinski M, Bick U, Zhang M, Schmidt RA, Wolverton DE, Comstock CE, Papaioannou J, Collins SA, Urbas AM, Vyborny CJ, Doi K: Computerized detection of breast lesions in digitized mammograms and results with a clinically-implemented intelligent workstation. In: Lemke HU, Vannier MW, and Inamura K (eds.), <u>CARS</u>

'97 Computer Assisted Radiology and Surgery. (Amsterdam: Elsevier
Science B.V.) pp. 325-330, 1997.

(24)   Giger ML, Nishikawa RM, Vyborny CJ, Schmidt RA, Wolverton DE,
Comstock C, Metz CE, Doi K: Development of methods for
computerassisted interpretations of digital mammograms for early breast
cancer detection. Proc. Era of Hope, Department of Defense Breast
Cancer Research Program Meeting, Vol. I, pgs. 83-84, 1997.

(25)   Krupinski EA, Nishikawa RM: Comparison of eye-position versus
computer identified microcalcification clusters on mammograms. Med
Phys 24: 17-23, 1997.

(26)   Nishikawa RM, Wolverton DE, Schmidt RA, Papaioannou J:
Radiologists' ability to discriminate computer-detected true and false
positives from an automated scheme for the detection of clustered
microcalcifications on digital mammograms. Proc SPIE 3036: 198-204,
1997.

(27)   Nishikawa RM, Giger ML, Jiang Y, Huo Z, Doi K, Schmidt RA,
Wolverton DE, Vyborny CJ: Automated Classification of Breast Lesions
on Digital Mammograms. In: Lemke HU, Vannier MW, and Inamura K
(eds.), CARS '97 Computer Assisted Radiology and Surgery.
(Amsterdam: Elsevier Science B.V.) pp. 325-330, 1997.

(28)   Nishikawa RM: The transfer of intelligence community and other
imaging technologies to improve women's health. In: Journal of the
American Medical Informatics Association, pg. 85, 1997.

(29)   Giger ML, Huo Z, Wolverton DE, Vyborny CJ, Moran C, Schmidt RA,
Al-Hallaq H, Nishikawa RM, Doi K: Computer-aided diagnosis of digital
mammographic and ultrasound images of breast mass lesions. In: Digital
Mammography (eds. N. Karssemeijer, M. Thijssen, J. Hendriks, L. van

Erning). Dordrecht (Kluwer Academic Publishers), pp. 143-148, 1998.

(30) Huo Z, Giger ML, Vyborny CJ, Wolverton DE, Schmidt RA, Doi K:
      Automated computerized classification of malignant and benign mass
      lesions on digitized mammograms. Acad Radiol 5: 155-168, 1998.

(31) Jiang Y, Nishikawa RM, Papaioannou JP. Requirement of
      microcalcification detection for computerized classification of malignant
      and benign clustered microcalcifications. Proc SPIE 3338: 313-317, 1998.

(32) Jiang Y, Nishikawa RM, Schmidt RA, Metz CE, Giger ML, Doi K:
      Benefits of computer-aided diagnosis (CAD) in mammographic diagnosis
      of malignant and benign clustered microcalcifications. In: Digital
      Mammography (eds. N. Karssemeijer, M. Thijssen, J. Hendriks, L. van
      Erning). Dordrecht (Kluwer Academic Publishers), pp. 215-220, 1998.

(33) Kupinski MA, Giger ML: Feature selection and classifiers for the
      computerized detection of mass lesions in digital mammography. Proc
      IEEE International Congress on Neural Networks 1997, pp. 1336-1339,
      1998.

(34) Kupinski MA, Giger ML: Automated seeded lesion segmentation on
      digital mammograms. IEEE Trans on Medical Imaging 17: 510-517,
      1998.

(35) Kupinski M, Giger ML: Investigation of regularized neural networks for
      the computerized detection of mass lesions in digital mammograms. Proc.
      EMBS'97, pp. 1336-1339, 1998.

(36) Nagel RH, Nishikawa RM, Doi K: Analysis of methods for reducing false
      positives in the automated detection of clustered microcalcifications in
      mammograms. Med Phys 25: 1502-1506, 1998.

(37) Nishikawa RM, Giger ML, Wolverton DE, Schmidt RA, Comstock CE,
      Papaioannou J, Collins SA, Doi K: Prospective testing of a clinical

43

mammography workstation for CAD: Analysis of the first 10,000 cases. In: Digital Mammography (eds. N. Karssemeijer, M. Thijssen, J. Hendriks, L. van Erning). Dordrecht (Kluwer Academic Publishers), pp. 401-406, 1998.

(38) Nishikawa RM, Yarusso LM: Variations in measured performance of CAD schemes due to database composition and scoring protocol. Proc SPIE 3338: 840-844, 1998.

(39) Nishikawa RM: Mammographic databases. Breast Disease 10: 137-150, 1998.

(40) Schmidt RA, Nishikawa RM, Jiang Y, Metz CE, Wolverton DE, Doi K, Giger ML, Cannon W: Can computers help radiologists decide who needs a breast biopsy? In: Karssemeijer N, et al. (eds.), Digital Mammography (Amsterdam: Kluwer), pgs. 401-406, 1998.

(41) Schmidt RA, Newstead GM, Linver MN, Eklund GW, Metz CE, Winkler MA, Nishikawa RM: Mammographic screening: Sensitivity of general radiologists. In: Karssemeijer N, et al. (eds.), Digital Mammography (Amsterdam: Kluwer), pgs. 383-388, 1998.

(42) Zhang W, Yoshida H, Nishikawa RM, Doi K: Optimally weighted wavelet transform based on supervised training for detection of microcalcifications in digital mammograms. Med Phys 25: 949-956, 1998.

(43) Anastasio MA, Kupinski MA, Nishikawa RM, Giger ML: Optimization of computer-aided diagnosis schemes using a multiobjective approach. Proc. 1998 IEEE Med Imaging Conference pgs. 1879-1883, 1999.

(44) Giger ML: Overview of computer-aided diagnosis in breast imaging. In: Computer Aided Diagnosis in Medical Imaging, Doi K, MacMahon H, Giger ML, Hoffmann KR (eds). (Elsevier, Amsterdam), pgs., 167-176,

1999.

(45)   Giger ML: "Computer-aided diagnosis". In: <u>AAPM/RSNA Categorical</u>
<u>Course in Diagnostic Radiology Physics: Physical Aspects of Breast</u>
<u>Imaging -- Current and Future Considerations</u>, Haus A. and Yaffe M.
(eds.) pp. 249-272, 1999.

(46)   Giger ML, Huo Z: Artificial neural networks in breast cancer diagnosis:
Merging of computer-extracted features from breast images. <u>Proc Of</u>
<u>Conference on Evolutionary Computing (CEC'99)</u>, pp. 1768-1769, 1999.

(47)   Huo Z, Giger ML: Robustness of a computerized scheme for the
classification of malignant and benign masses on digitized mammograms.
In: <u>Computer Aided Diagnosis in Medical Imaging</u>, Doi K, MacMahon H,
Giger ML, Hoffmann KR (eds). (Elsevier, Amsterdam), pgs. 277-280,
1999.

(48)   Huo Z, Giger ML, Metz CE: Effect of dominant features on neural
network performance in the classification of mammographic lesions. <u>Phys</u>
<u>Med Biol</u> 44: 2579-2595, 1999.

(49)   Jiang Y, Nishikawa RM, Schmidt RA, Metz CE, Giger ML, Doi K:
Improving breast cancer diagnosis with computer-aided diagnosis. <u>Acad</u>
<u>Radiol</u> 6: 22-33, 1999.

(50)   Jiang Y, Nishikawa RM, Schmidt RA, Metz CE, Giger ML, Doi K:
Improvement in radiologists' diagnosis of malignant and benign clustered
microcalcifications by the use of computer-aided diagnosis (CAD). In:
<u>Computer Aided Diagnosis in Medical Imaging</u>, Doi K, MacMahon H,
Giger ML, Hoffmann KR (eds). (Elsevier, Amsterdam), pgs. 233-236,
1999.

(51)   Jiang Y, Nishikawa RM: Radiologists' ability of using computer-aided
diagnosis (CAD) to improve breast biopsy recommendations. <u>Proc SPIE</u>

45

3663: 56-60, 1999.

(52)   Kupinski MA, Giger ML: Feature selection with limited datasets. <u>Med Phys</u> 26: 2176-2182, 1999.

(53)   Nishikawa RM: Computer-aided diagnosis in digital mammography. <u>Diagnostic Imaging</u> 75: 47-51, 1999.

(54)   Nishikawa RM, Giger ML, Schmidt RA, Wolverton DE, Doi K: Prospective testing of a clinical CAD workstation for the detection of breast lesions on mammograms. In: <u>Computer Aided Diagnosis in Medical Imaging</u>, Doi K, MacMahon H, Giger ML, Hoffmann KR (eds). (Elsevier, Amsterdam), pgs. 209-214, 1999.

(55)   Yoshida H, Anastasio M, Nagel R, Nishikawa RM, Doi K: Computer-aided diagnosis for detection of clustered microcalcifications in mammograms: Automated optimization of performance based on genetic algorithm. In: <u>Computer Aided Diagnosis in Medical Imaging</u>, Doi K, MacMahon H, Giger ML, Hoffmann KR (eds). (Elsevier, Amsterdam), pgs. 247-252, 1999.

(56)   Edwards DC, Kupinski MA, Nagel RH, Nishikawa RM, Papaioannou J: Using a Bayesian Neural Network to Optimally Eliminate False-Positive Microcalcification Detections in a CAD Scheme. In: <u>Digital Mammography 2000</u>, Yaffe MJ, (ed). (Medical Physics Publishing, Madison WI) pgs. 168-173, 2000.

(57)   Giger ML, Huo Z, Kupinski MA, Vyborny CJ: Computer-aided diagnosis in mammography. In: <u>The Handbook of Medical Imaging, volume 2. Medical Imaging Processing and Analysis</u>, Fitzpatrick JM, Sonka M, (eds.) SPIE, pgs. 915-1004, 2000.

(58)   Giger ML, Huo Z, Lan L, Vyborny CJ: Intelligent search workstation for computer-aided diagnosis. <u>Proc of Computer Assisted Radiology and</u>

Surgery (CARS 2000), pgs. 822-827, 2000.

(59)   Giger ML: Computer-aided diagnosis of breast lesions in medical images. Computing in Science and Engineering Sept/Oct: 39-45, 2000.

(60)   Huo Z, Giger ML: Evaluation of a computer segmentation method based on performances of an automated classification method. Proc SPIE 2000 3981: 16-21, 2000.

(61)   Huo Z, Giger ML, Vyborny CJ, Wolverton DE, Metz CE: Computerized classification of benign and malignant masses on digitized mammograms: A Robustness study. Acad Radiol 7:1077-1084, 2000

(62)   Jiang Y, Nishikawa RM, Schmidt RA, Metz CE, Doi K: Relative gains in diagnostic accuracy between computer-aided diagnosis and independent double reading. Proc SPIE 3981: 10-15, 2000.

(63)   Jiang Y. Classification of breast lesions in mammograms. In: Bankman I Eds., Handbook of Medical Imaging Processing and Analysis, New York: Academic Press, pgs. 341-357, 2000.

(64)   Kupinski MA, Giger ML: Multiobjective genetic optimization of diagnostic classifiers used in the computerized detection of mass lesions in mammography. Proc SPIE 2000 3979: 40-45, 2000.

(65)   Kupinski MA, Giger ML: A comparison of Bayesian ANN and multiobjective classifier training using limited datasets. Proc CARS 2000, 752-757, 2000.

(66)   Nishikawa RM, Giger ML, Schmidt RA, Vyborny CJ, Bick J, Doi K: Prospective computer analysis of cancers missed on screening clinical. In: Digital Mammography 2000, Yaffe MJ, (ed). (Medical Physics Publishing, Madison WI) 493-498, 2000.

(67)   Valverde FL, Muñoz J, Nishikawa RM, Doi K: Elimination of calcified false positives in detection of microcalcifications in mammograms using

Hough transform. In: <u>Digital Mammography 2000</u>, Yaffe MJ, (ed).
(Medical Physics Publishing, Madison WI), pgs. 383-389, 2000.

(68) Vyborny CJ, Doi T, O'Shaughnessy KF, Harlan HM, Schneider AC, Stein
AA: Importance of spiculation in the computer detection of breast cancer.
<u>Radiology</u> 215: 703-707, 2000.

(69) Vyborny CJ, Giger ML, Nishikawa, RM: Computer-aided detection and
diagnosis. <u>Radiologic Clinics of North America</u> 38: 725-740, 2000.

(70) Edwards DC, Kupinski MA, Nagel RH, Nishikawa RM, Papaioannou J:
Using a Bayesian neural network to optimally eliminate false-positive
microcalcification detections in a CAD scheme. In: <u>Digital</u>
<u>Mammography 2000</u>, Yaffe MJ, et al. (eds). (Medical Physics Publishing,
Madison WI) pp. 168-173, 2001.

(71) Edwards DC, Papaioannou J, Jiang Y, Kupinski MA, Nishikawa RM:
Eliminating false-positive microcalcification clusters in a mammography
CAD scheme using a Bayesian neural network. <u>Proc SPIE</u> 4322: 1954-
1960, 2001.

(72) Giger ML, Maloney M, Huo Z, Vyborny CJ, Kupinski M, Venta L:
Computerized classification of lesions on digital mammography. <u>Digital</u>
<u>Mammography 2000, Proc. 5th International Workshop on Digital</u>
<u>Mammography, Medical Physics Publishing</u>, Wisconsin, pp. 189-1902,
2001.

(73) Huo Z, Giger ML, Vyborny CJ: Computerized analysis of multiple-
mammographic views: Potential usefulness of special view mammograms
in computer-aided diagnosis. <u>IEEE Transactions on Medical Imaging</u> 20:
1285-1292, 2001.

(74) Jiang Y, Nishikawa RM, Papaioannou J: Dependence of computer

classification of clustered microcalcifications on the correct detection of microcalcifications. Med Phys 28: 1949-1957, 2001.

(75)   Jiang Y, Nishikawa RM, Schmidt RA, Toledano AY, Doi K. The potential of computer-aided diagnosis (CAD) to reduce variability in radiologists' interpretation of mammograms containing microcalcifications. Radiology, 220: 787-794, 2001.

(76)   Jiang Y, Nishikawa RM, Venta L, Maloney MM, Giger ML: Computer classification of malignant and benign microcalcifications in small-field digital mammograms. Digital Mammography 2000, Proc. 5th International Workshop on Digital Mammography, Medical Physics Publishing, Madison, WI, pgs. 237-242, 2001.

(77)   Nishikawa RM, Giger ML, Schmidt RA, Vyborny CJ, Bick J, Doi K: Prospective computer analysis of cancers missed on screening mammography. Digital Mammography 2000, Proc. 5th International Workshop on Digital Mammography, Medical Physics Publishing, Madison WI pp. 493-498, 2001.

(78)   Nishikawa RM, Giger ML, Schmidt RA, Papaioannou J: Can computer-aided diagnosis (CAD) help radiologists find mammographically missed screening cancers? Proc. SPIE 4324: 56-63, 2001.

(79)   Huo Z, Giger ML, Vyborny CJ, Metz CE: Effectiveness of CAD in the diagnosis of breast cancer: An observer study on an independent database of mammograms. Radiology, 224: 560-568, 2002.

(80)   Huo Z, Giger ML: Effect of case-mix on feature selection in the computerized classification of mammographic lesions. Proc. SPIE, 4684: 762-767, 2002.

(81)   Nishikawa RM, Salfity MF, Jiang Y, Papaioannou J. Improving the automated classification of clustered calcifications on mammograms

49

through the improved detection of individual calcifications. <u>Proc SPIE</u>, 4684: 1339-1345, 2002.

(82) Vyborny C, Kukec C, Jiang Y, Doi K: Experience with computer-aided detection in a low-volume mammography clinic. <u>Proc 6th International Workshop on Digital Mammography</u>, (in press), 2002.

7. <u>LIST OF PERSONNEL</u>

Kunio Doi, Ph.D., Professor

Charles E. Metz, Ph.D., Professor

Robert A. Schmidt, M.D., Professor

Robert M. Nishikawa, Ph.D., Associate Professor

Dulcy E. Wolverton, M.D., Assistant Professor

Christopher E. Comstock, M.D., Assistant Professor

Hiroyuki Yoshida, Ph.D., Assistant Professor

Qiang Li, Ph.D., Research Associate (Assistant Professor)

Hiroyuki Abe, Ph.D., Research Associate

Martin Fiebich, Ph.D., Research Associate

Shigehiko Katsuragawa, Ph.D., Research Associate

Yoshitaka Masutani, Ph.D., Research Associate

Kenji Suzuki, Ph.D., Research Associate

Roger M. Engelmann, B.S., Computer Scientist

Chun-Wai Chan, B.S., Computer Scientist

Michael D. Carlin, R.T. Technologist

8. REFERENCES

(1) Wingo PA, Tong T, Bolden S: Cancer Statistics, 1995. CA 45: 8-30, 1995.

(2) Lissner J, Kessler M, Anhalt G: Developments in methods for early detection of breast cancer. In: Early Breast Cancer by J. Zander and J. Baltzer, eds. (Springer-Verlag, Berlin 1985), pp. 93-123.

(3) Bassett LW, Gold RH: Breast Cancer Detection: Mammography and Other Methods in Breast Imaging, (Grune and Stratton, New York, 1987).

(4) Andersson I, What can we learn from interval carcinomas? Recent Results in Cancer Research 90: 161-163, 1984.

(5) Martin JE, Moskowitz M, Milbrath JR: Breast cancers missed by mammography. Am. J. Roentgenol. 132: 737, 1979.

(6) Holland T, Mrvunac M, Hendrirks JHCL, Bekker BV: So-called interval cancers of the breast. Pathologic and radiographic analysis. Cancer 49: 2527-2533, 1982.

(7) Buchanan JR, Spratt JS, Heuser LS: Tumor growth, doubling times, and the inability of the radiologist to diagnosis certain cancers. Radiologic Clinics of North America 21: 115-126, 1983.

(8) Chan H-P, Doi K, Vyborny CJ, Schmidt RA, Metz CE, et al.: Improvement in radiologists' detection of clustered microcalcifications on mammograms: The potential of computer-aided diagnosis. Invest Radiol 25: 1102-1110, 1990.

(9) Astley S, Hutt I, Adamson S, Miller P, Rose P: Automation in mammography: computer vision and human perception. Proc. SPIE 1905: 716-730, 1993.

(10) Doi K, Giger ML, Nishikawa RM, Hoffmann KR, MacMahon H, Schmidt RA, Chua KG: Digital radiography: A useful clinical tool for computer-aided diagnosis by quantitative analysis of radiographic images. Acta Radiologica 34: 426-439, 1993.

(11) MacMahon H, Doi K: Digital chest radiography. Clinics in Chest Medicine 12: 19-32, 1991.

(12) Giger ML, Doi K, MacMahon H: Image feature analysis and computer-aided diagnosis in digital radiography: 3. Automated detection of nodules in peripheral lung field. Med Phys 15: 158-166, 1988.

(13) Matsumoto T, Yoshimura H, Doi K, Giger ML, et al.: Image feature analysis of false positives produced by an automated computerized scheme for the detection of lung nodules in digital chest radiographs. Invest Radiol 27: 587-597, 1992.

(14) Katsuragawa S, Doi K, MacMahon H, et al.: Quantitative analysis of lung texture in the ILO pneumoconiosis standard radiographs. RadioGraphics 10: 257-269, 1990.

(15) Chen X, Doi K, Katsuragawa S, MacMahon H: Automated selection of regions of interest for quantitative analysis of lung textures in digital chest radiographs. Med Phys 20: 975-982, 1993.

(16) Nakamori N, Doi K, Sabeti V, MacMahon H: Image feature analysis and computer-aided diagnosis in digital radiography: Automated analysis of sizes of heart and lung in digital chest images. Med Phys 17: 342-350, 1990.

(17) Nakamori N, Doi K, MacMahon H, Sasaki Y, Montner S: Effect of heart size parameters computed from digital chest radiographs on detection of cardiomegaly: Potential usefulness for computer-aided diagnosis. Invest Radiol 26: 546-550, 1991.

(18) Sanada S, Doi K, MacMahon H: Image feature analysis and computer-aided diagnosis in digital radiography: Automated detection of pneumothorax in chest images. Med Phys 19: 1153-1160, 1992.

(19) Hoffmann KR, Doi K: Correspondence regarding "Determination of instantaneous and average blood flow rates from digital angiograms using distance-density curves." Invest Radiol 27: 274, 1992.

(20) Alperin N, Hoffmann KR, Doi K, Chua KG: Automated analysis of coronary lesions from cineangiograms using vessel tracking and iterative deconvolution

techniques. <u>Proc. of Computers in Cardiology</u>, Jerusalem, Israel, Sept. 19-22, IEEE, pp. 153-156, 1990.

(21) Caligiuri P, Giger ML, Favus M, Jia H, Doi K, Dixon LB: Computerized radiographic analysis of osteoporosis. <u>Radiology</u> 186: 471-474, 1993.

(22) Giger ML, Vyborny CJ: CAD in mammography: rationale, methods and possible scenarios. <u>Diagnostic Imaging</u>, June, pgs. 98-113, 1993.

(23) Giger ML: "Future of Breast Imaging. Computer-Aided Diagnosis". Book chapter in: <u>AAPM/RSNA Categorical Course on the Technical Aspects of Breast Imaging</u>, (eds.) Haus A. and Yaffe M. pgs. 257-270, 1992.

(24) Chan HP, Doi K, Galhotra S, Vyborny CJ, MacMahon H, Jokich PM: Image feature analysis and computer-aided diagnosis in digital radiography. 1. Automated detection of microcalcifications in mammography. <u>Med Phys</u> 14: 538-548, 1987.

(25) Chan HP, Doi K, Vyborny CJ, Lam KL, Schmidt RA: Computer-aided detection of microcalcifications in mammograms: Methodology and preliminary clinical study. <u>Invest Radiol</u> 23: 664-671, 1988.

(26) Wu Y, Doi K, Giger ML, Nishikawa RM: Computerized detection of clustered microcalcifications in digital mammograms: Application of artificial neural networks. <u>Med Phys</u> 19: 555-560, 1992.

(27) Zhang W, Itoh K, Tanida J, Ichioka Y: Parallel distributed processing model with local space-invariant interconnections and its optical architecture. <u>Applied Optics</u> 29: 4790-4797, 1990.

(28) Zhang W, Doi K, Giger ML, Wu Y, Nishikawa RM, Schmidt RA: Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network. <u>Med Phys</u> 21:517-524, 1994.

(29) Yin FF, Giger ML, Doi K, Metz CE, Vyborny CJ, Schmidt RA: Computerized detection of masses in digital mammograms: Analysis of bilateral subtraction images. <u>Med Phys</u> 18: 955-963, 1991.

(30) Yin FF, Giger ML, Vyborny CJ, Doi K, Schmidt RA: Comparison of bilateral-subtraction and single-image processing techniques in the computerized detection of mammographic masses. Invest Radiol 28: 473-481, 1993.

(31) Yin FF, Giger ML, Doi K, Vyborny CJ, Schmidt RA: Computerized detection of masses in digital mammograms: Investigation of feature-analysis techniques. J. Dig. Img 7:18-26, 1994.

(32) Yin FF, Giger ML, Doi K, Vyborny CJ, Schmidt RA: Computerized detection of masses in digital mammograms: Automated alignment of breast images and its effect on bilateral-subtraction technique. Med Phys 21:445-452, 1994.

(33) Nishikawa RM, Giger ML, Doi K, Metz CE, Yin F-F, Vyborny CJ, Schmidt RA: Effect of case selection of the performance of computer-aided detection schemes. Med Phys 21:265-269, 1994.

(34) Nishikawa RM, Haldemann RC, Papaioannou J, Giger ML, Lu P, Schmidt RA, Wolverton DE, Bick U, Doi K: Initial experience with a prototype clinical "intelligent" mammography workstation for computer-aided diagnosis. Proc SPIE 2434: 65-71, 1995.

(35) Fukushima K, Miyake S, Ito T: "Neocognitron: A neural model for a mechanism of visual pattern recognition," IEEE Trans. Systems Man and Cybernetics SMC 13: 826-843, 1983.

(36) Fukushima K: "A neural network for visual pattern recognition," Computer 21: 65-76, 1988.

(37) Schmidt RA, Nishikawa RM, Schreibman K, Giger ML, Doi K, Papaioannou J, Lu P, Stucka J, Birkhahn G: Computer detection of lesions missed by mammography. Proc 2nd Int Workshop on Digital Mammography, Edited by Gale AG, et al, published by Elsevier, pp. 289-294, 1994.

(38) Jiang Y, Metz CE, Nishikawa RM: An ROC partial area index for highly sensitive diagnostic tests. Radiology 201: 745-750, 1996.

(39) Huo Z, Giger ML, Vyborny CJ, et al.: Analysis of speculation in the computerized classification of mammographic masses. <u>Med Phys</u> 22: 1569-1579, 1995.

(40) Kupinski MA, Giger ML: Investigation of regularized neural networks for the computerized detection of mass lesions in digital mammograms. Proc. EMBS'97. pp. 1336-1339, 1998.

(41) Kupinski MA, Giger ML: Automated seeded lesions segmentation on digital mammograms. IEEE Trans. Medical Imaging 17: 510-517, 1998.

(42) Kupinski MA, Giger ML: Feature selection with limited datasets. Medical Physics, 26: 2176-2182, 1999.

(43) Huo Z, Giger ML, Metz CE: Effect of dominant features on neural network performance in the classification of mammographic lesions. Phys. Med. Biol., 44: 2579-2595, 1999

(44) Jiang Y, Nishikawa RM, Venta LL, Maloney MM, Giger ML. Computer classification of malignant and benign microcalcifications in small-field digital mammograms. In Digital Mammography (Yaffe MJ, Eds.). Madison, WI: Medical Physics Publishing, pp. 215-220, 2000.

(45) Jiang Y. Classification of malignant and benign breast lesions in mammograms. In <u>Handbook of Medical Image Processing</u> (Bankman I, Eds.). New York: Academic Press, pp. 341-357, 2000.

(46) Jiang Y, Nishikawa, RM, Schmidt RA, Toledano AY, Doi K. The potential of computer-aided diagnosis (CAD) to reduce variability in radiologists' interpretation of mammograms. <u>Radiology</u> 220:787-794, 2001.

(47) Jiang Y, Nishikawa RM, Papaioannou J. Dependence of computer classification of clustered microcalcifications as malignant or benign on the correct detection of microcalcifications. <u>Medical Physics</u> 28:1949-1957, 2001.

(48) Beiden SV, Wagner RF, Campbell G, Metz CE, Jiang Y. Components-of-variance models for random-effects ROC analysis: The case of unequal variance structures across modalities. Academic Radiology 8:605-615, 2001.

(49) Jiang Y, Metz CE. An optimal method for combining two correlated diagnostic assessments with application to computer-aided diagnosis. Proc SPIE 4324:177-183, 2001.

(50) Huo Z, Giger ML, Vyborny CJ, Metz CE: Effectiveness of CAD in the diagnosis of breast cancer: An observer study on an independent database of mammograms. Radiology, 224: 560-568, 2002.

(51) Nishikawa RM, Giger ML, Schmidt RA, Vyborny CJ, Bick J, Doi K: Prospective computer analysis of cancers missed on screening clinical. In: Digital Mammography 2000, Yaffe MJ, et al. (eds). (Medical Physics Publishing, Madison WI) pp. 493-498, 2000.

(52) Nishikawa RM, Giger ML, Schmidt RA, Papaioannou J: Can computer-aided diagnosis (CAD) help radiologists find mammographically missed screening cancers? Proc. SPIE, 4324: 56-63, 2001.